Computer Engineering and Intelligent Systems
ISSN 2222-1719 (Paper) ISSN 2222-2863 (Online)
Vol.6, No.10, 2015
www.iiste.org

Fraud Detection in Telecommunications Industry: Bridging the Gap with Random Rough Subspace Based Neural Network Ensemble Method

Iyabo Awoyelu Adenike Adebomi Adekemi Amoo Rachael Adebisi Charles Mabude
Department of Computer Science and Engineering, Obafemi Awolowo University, Ile-Ife, Nigeria

Abstract
Fraud has been very common in the society and it affects private enterprises as well as public entities. Telecommunication companies worldwide suffer from customers who use the provided services without paying. There are also different types of telecommunication fraud such as subscription fraud, clip on fraud, call forwarding, cloning fraud, roaming fraud and calling card fraud. Thus, detection and prevention of these frauds are the main targets of the telecommunication industry. This paper addresses the various techniques of detecting fraud, giving the limitations of each technique and proposes random rough subspace-based neural network ensemble method for effective fraud detection.
Keywords: Fraud, Fraud detection, Random rough subspace, Neural network, Telecommunications

1. Introduction

Fraud is considered as attractive from the fraudster's point of view, since detection risk is low, no special equipment is needed, and the product in question is easily converted to cash. Telecommunication companies have a long history of fighting fraud but the imaginative methods employed by fraudsters call for continuous improvements on the solutions deployed to check fraud. In spite of the fact that many telecommunication companies worldwide have lost significant amounts of money due to fraudulent activity on their networks, large numbers of operators are still not addressing this critical issue. In many cases, they even feel that fraud does not exist. Even though one wishes that this should be the case, this is never true. Moreover, losses due to fraud are often swept under the carpet as bad debts.

Also, there is the belief that networks based on digital technologies are secure. Innovative fraudsters have managed to find simple, non-technical ways to continue their notorious activities even in technically advanced digital networks. Fraud will never be fought unless it is acknowledged. Unfortunately, many operators, especially in developing markets are on an overdrive to attract more and more subscribers to the network. Some operators are just starting to realize that many subscribers are only fraud generating and not revenue generating. Understanding the nature of fraud is the first step to reduce revenue leakage. Although prevention technologies are the ways of reducing fraud, fraudsters when given time will usually find ways to circumvent such measures. Methodologies for the detection of fraud are therefore essential if one is to catch fraudsters, once fraud prevention has failed.

Detecting fraud is hard, so it is not surprising that many fraud systems have serious limitations. Different systems may be needed for different kinds of fraud with each system having different procedures, different parameters to tune, different database interface, different case management tools and features. This paper is concerned with detection of fraud in telecommunications industry using random rough subspace-based neural network ensemble method.

The remaining part of this paper is organized as follows. Section 2 discusses Bayesian network technique. Section 3 explains Rough set Ensemble method, how it can be used for detection of fraud in telecommunications industry and Section 4 concludes the paper.

2.0 Fraud Detection Techniques

Fraud detection methods can be supervised or unsupervised (Hilas and Sahalos, 2007). Supervised methods are those where samples of both normal and fraudulent behaviour are used to construct models, which enable the system to assign new observations to one of the two classes. One must have data of both classes and should also be sure about the true class in which original observations belong to. Moreover, this method can only identify known fraudulent activities.

Unsupervised methods simply seek those observations that are dissimilar from the norm. They usually deal with outlier or any other extreme data detection. Research in telecommunication fraud detection is mainly motivated by fraudulent activities in mobile technologies (Patidar and Sharma, 2011).

2.1 Bayesian Network Technique

According to Taniguchi *et al.* (2000), there are no deterministic rules which allow someone to identify a subscriber as a fraudster. One may at best formulate one's degree of belief in fraudulent behaviour. Graphical models such as

Bayesian networks supply a general framework for dealing with uncertainty in a probabilistic setting and thus they are well suited to tackle the problem of fraud detection. Every graph of a Bayesian network codes a class of probability distributions. The nodes of that graph comply with the variables of the problem domain. Arrows between nodes denote allowed (causal) relations between the variables. These dependencies are quantified by conditional distributions for every node given its parents. Once a Bayesian network is set up, one can infer probabilities for unknown variables by inserting evidence in the network and propagating the evidence through the network using propagation rules (Taniguchi *et al.*, 2000). For the purpose of fraud detection, Taniguchi *et al.* (2000) construct two Bayesian networks to describe the behaviour of mobile phone subscribers. First, a Bayesian network was constructed to model behaviour under the assumption that the subscriber is fraudulent (F) and another model under the assumption the subscriber is a legitimate user (NF).This is as shown in Figure 1.

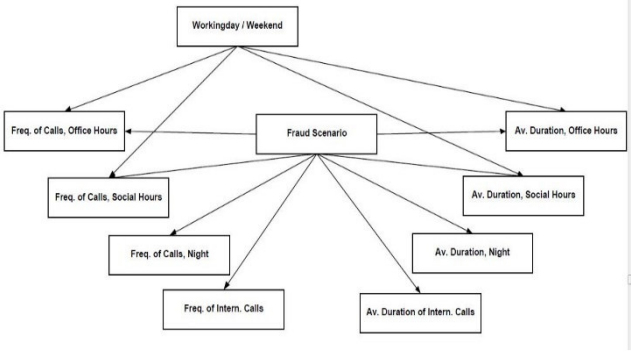

Figure 1. Bayesian Network in Fraud Detection (Taniguchi *et al*., 2000)

The `fraud net' is set up by using expert knowledge. The user net is set up by using data from non-fraudulent subscribers.

On the other hand, Nelson (2009) uses a Bayesian network as a graphical model that encodes probabilistic relationship among variables. Every graph of a Bayesian network codes a class of probability distributions. The nodes of that graph comply with the variables of the problem domain. Arrows between nodes denote allowed (casual) relations between the variables. In this technique, a relationship is set up between each piece of knowledge and associates a probability that gives piece of knowledge, how much that particular piece of knowledge influences the event B, the event being in this case is the probability of the customer being fraud. For example, given that the average call duration is X and most calls occur in the evening, is the customer fraudulent? This technique uses two beliefs networks. The first network is modeled with the relationship between knowledge being established based on the previous fraud that has been detected. The second is a network that is automatically generated from all the clear (non-fraudulent) data in the network and a network is normally created for each customer class. The data for each customer are then passed through both networks and the results from both networks are considered on containing a belief of how fraudulent a customer is and the second belief of how clear a customer is. This technique is not sufficient since if some important relationships are missed out during inferring knowledge in the system, the system will not respond properly. Also if the customer is perpetrating a new type of fraud that has never been modeled before, the system will still not respond properly.

2.2 Distance-based Method

An outlier is an observation that deviates so much from other observations as to arouse suspicion, or the set of data points that are considerably different from the remainder of the data (Nelson, 2009). One of the outlier applications, distance-based method was originally proposed by Rajani and Padmavathamma (2012). This notion is further extended based on the distance of a point from its *kth* nearest neighbor. Alternatively, the outlier factor of each data point is computed as the sum of distances from its *k* nearest neighbors. The drawbacks of distance-based methods are that it is hard to find clusters (collection of data objects that are similar to one another) and it is hard to specify the number of clusters. Hence it may also give rise to many false positive alerts.

2.3 Time-series Analysis

In time-series analysis, outliers are found using peer group analysis (PGA). Peer group analysis (PGA) is the term used to describe the analysis of the time evolution of a given object (the target) relative to other objects that have been identified as initially similar to the target in some senses (the peer group) (Serrano *et al.*, 2010). This is an unsupervised technique for fraud detection whereby expected patterns of behaviour of similar objects are characterized in terms of the behaviour of similar objects and any difference in evolution between the expected pattern and the target is detected. This is not suitable for detecting subscription fraud and bad debts as may lead to many false positive alerts.

2.4 Rule-based Approach to Fraud Detection

In rule-based methods, the usage data are verified against specific rules of the form: If certain conditions, Then consequence. Based on the knowledge obtained from the human experts in telecommunication fraud, a set of rules can be created that can match certain aspects of a customer profile with a set threshold. These rules may be absolute, based on simple thresholds which may or may not be customer dependent, or differential based on observed statistical anomalies, the identification of which can be based on customer profile, time of the day or other factors (Nelson, 2009).

Rule-based methods have some limitations. They require knowledge of the exact parameters of fraud. In addition, since there are seemingly unlimited methods to defraud, it would imply that the rule set required to capture the fraudsters would also need to be sufficiently large. This is not feasible considering each check may take a finite period of time, and the larger the rules, the longer the checks will take. And, that assumes to give possibly a little gain in fraud detection.

Also, rule systems are not dynamic in their nature (Verrelst *et al.*, 2001). It is rarely practical to access or analyze all call detail records for an account every time it is evaluated for fraud (Gopal and Meher, 2007). Hence, a common approach is to reduce the call records for an account to several statistics that are computed each period. The summaries that are monitored for fraud may be defined by subject matter experts and thresholds may be chosen by trial and error. Decision trees or machine learning algorithms may be applied to training set of summarized account data to determine good threshold rules. Thresholds have some disadvantages. Although, they may need to vary with type of account, type of call, and time of the day to be sensitive to fraud without setting of too many false alarms for legitimate accounts (Gopal and Meher, 2007).

In Advanced Security for Personal Communication Technologies (ASPeCT), several approaches are taken to identify fraudulent behaviours (Verrelst *et al.*, 2001). In the rule-based approach, both the absolute and differential usage is verified against certain rules. This approach works best with user profiles containing explicit information, where fraud criteria given as rules can be referred to. User profiles are maintained for the directory number of the calling party (A-number), for the directory number of the called party (B-number) and also for the cells used to make/receive the calls. A-number profiles represent user behaviour and are useful for the detection of most types of fraud, while B-number profiles point to hot destinations and thus allow the detection of frauds based upon call forwarding. All deviations from normal user behaviour resulting from the different analysing processes are collected and alarms will finally be raised if the results in combination fulfill given alarm criteria. The implementation of this solution is based on an existing rule-based tool for audit trail analysis Protocol Data Analysis Tool (PDAT) (Verrelst *et al.*, 2001). Intrusion detection and mobile fraud detection are quite similar problem fields; the flexibility and broad applicability of PDAT are promising for using this tool for mobile fraud detection. The main tasks are the introduction of user profiles stored in a database and the realization of a new protocol that allows PDAT to understand both user profile as well as Toll Ticket formats. Once established, PDAT provides a comprehensive infrastructure based on a GUI for showing alarms and for editing alarm criteria during runtime. The new architecture is depicted in Figure 2.

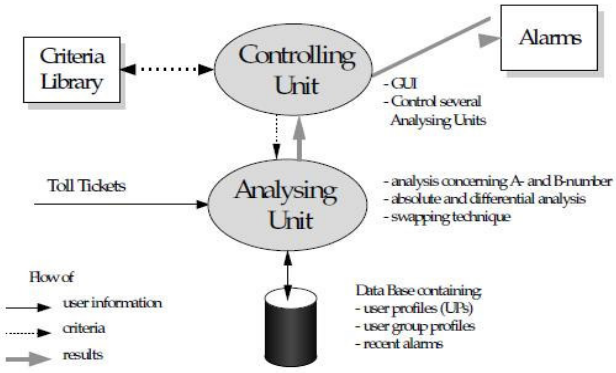

Figure 2. Architecture of Rule-based Fraud Detection Tool (Verrelst *et al.*, 2001)

2.5 Neural Network Based Approach to Fraud Detection

Another approach to identify fraudulent behaviour uses neural networks. An artificial neural network consists of a collection of processing elements that are highly interconnected and transforms a set of inputs to a set of desired outputs. The result of the transformation is determined by the characteristics of the elements and the weights associated with the interconnections among them. The multiplicity and heterogeneity of the fraud scenarios require the use of intelligent detection systems. The fraud detection engine has to be flexible enough to cope with the diversity of fraud. It should also be adaptive in order to face new fraud scenarios, since fraudsters are likely to

develop new forms of fraud once old attacks become impractical. Further, frauds appear in the billing system as abnormal usage patterns in the Toll Ticket records of one or more users. The function of the fraud detection engine is to recognize such patterns and produce the necessary alarms. High flexibility and adaptability for a pattern recognition problem directly point to neural networks as a potential solution.

Neural networks are systems of elementary decision units that can be adapted by training in order to recognize and classify arbitrary patterns. The interaction of a high number of elementary units makes it possible to learn arbitrarily complex tasks (Taniguchi *et al.*, 2000; Verrelst *et al.*, 2001). As a closely related application, neural networks are now routinely used for the detection of credit card fraud. There are two main forms of learning in neural networks namely unsupervised learning and supervised learning. In supervised learning, the patterns have to be a *priori* labeled as belonging to some class (Taniguchi *et al.*, 2000). During learning, the network tries to adapt its units so that it produces the correct label at its output for each training pattern. Once training is finished, the units are frozen and when a new pattern is presented, it is classified according to the output produced by the network. In unsupervised learning, the system is allowed to find patterns or clusters in the data in the hope that these clusters will be useful or meaningful in some way, either directly or indirectly (Taniguchi *et al.*, 2000).

2.6 B-number Analysis Tool

The B-number analysis tool monitors the destination countries of calls on a per subscriber basis. The destinations of calls (the B-Number) are weighted differently so that well known destinations for fraudulent calls can be given a special attention. The profile is maintained as a probability distribution of the call destination for the Current Usage Profile (CUP) and Usage Profile History (UPH). The fraud engine takes the B-number profile record consisting of the CUP and UPH as input and calculates a modified distance over all the entries of the profile record (Nelson, 2009).

2.7 Agent-based Knowledge Discovery in Data

Telecommunication fraud is a problem that affects operators all around the world. Operators know that fraud cannot be completely eradicated. The solution to deal with this problem is to minimize the damages and cut down losses by detecting fraud situations as early as possible. In Sanver and Karahoca (2008), the fraud problem was analyzed and a new approach to the problem was designed. This new approach, based on the profiling and Knowledge Discovery in Data (KDD) techniques, supported in Multi-agent System (MAS), does not replace the existing fraud detection systems; it uses them and their results to provide operators with new fraud detection methods and new knowledge.

2.8 Multi-agent system

The main purpose of the Multiagent Systems (MAS) is the study, construction and application of multi-agent systems, that is, systems in which several interacting, intelligent agents pursue some set of goals and/or perform some set of tasks (Rosas and Analide, 2009).

The MAS implemented to support the solution proposal is a closed MAS, where the architecture design is static, with all the agents and functionalities predefined. In this closed MAS, the agents communicate using a common language, each agent is developed as an expert in his functionality and they all work and cooperate together in order to achieve a main goal. The coordination of the MAS is cooperative (Rosas and Analide, 2009): the agents do not compete; they cooperate in order to achieve a main objective. The organization is flat (Rosas and Analide, 2009): each agent is an expert in an area, there is no agent commanding other agents and all agents have the same importance and status. The communication between the agents is direct (Rosas and Analide, 2009): there is no agent or middleware between two agents communicating with each other, they communicate directly. Figure 3 illustrates the MAS architecture and the process flow between the MAS and the fraud solution. As for the architecture, the MAS is composed of three agents (Rosas and Analide, 2009) namely the Profiling agent, the Detecting agent and the KDD agent.

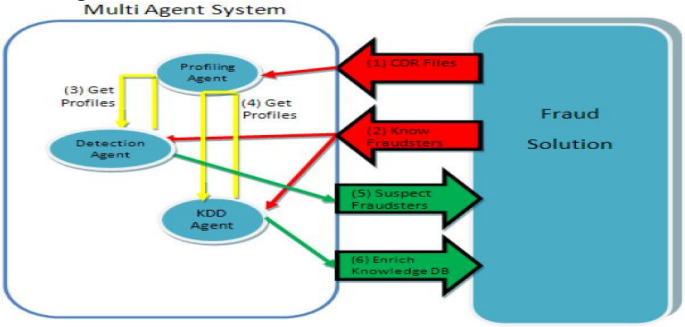

Figure 3. MAS - Architecture and Process Flow (Rosas and Analide, 2009)

The Profiling Agent is responsible for integrating the CDR profiles and building the profiles (with identity and behaviour features) for the subscribers. The Detection Agent is responsible for detecting new fraud suspects and the KDD Agent is responsible for processing the profiles of known fraudsters and enriching the knowledge database.

As for the process flow between the MAS and the fraud solution, the flow of information is represented by enumerated arrows (Rosas and Analide, 2009). In the first phase, the fraud solution redirects all the CDR files contents for the MAS (step 1 in Figure 3), where the profiling agent uses this content to build the profiles for the subscribers. In this phase, the profiling agent should extract from the CDR files contents the necessary information to in order to enrich the identity and behaviour features of profiles. In a second stage, when the fraud solution detects a new fraudster, based on some of the methods previously explained, the fraud solution indicates the MAS that a new fraudster is detected (step 2 in Figure 3). Then, this information is used by two agents with different purposes (Rosas and Analide, 2009):

- The detection agent will use this information in order to retrieve from the profiling agent, the profile of the fraudster and use the profile identity features to detect if the same subscriber tries to re-enter the operator network and the profile behaviour features detect other subscribers that have a similar behaviour; the fraud solution is then warned of the suspects that this agent detects (step 3 in Figure 3).
- The KDD agent will use this information in order to retrieve from the profiling agent, the profile of the fraudster and use the profile behaviour features and enrich a knowledge database containing all the detected fraudsters profiles (only the behaviour features) in order to try to retrieve significant information (patterns, similar behaviours) from this database. The results of this enrichment are then passed to the fraud solution (step 4 in Figure 3), so that the fraud analysts have access to this information.

2.9 User profiling

Profiling is an auxiliary technique for criminal investigation (Rosas and Analide, 2009). It fits in the Forensic Psychology domain. Profiling consists in a process of individual features inference, usually individuals responsible for criminal actions. The profiling technique should be used as an extension of the criminal analysis, elaborating criminal profiles based on previous work. The main idea to retain is: profiling complements previous work, it does not replace it. The main idea behind user profiling is that past behaviour of a user can be accumulated in order to construct a profile or a "user dictionary" of what might be the expected values of the user's behaviour (Hilas and Sahalos, 2007). This profile contains single numerical summaries of some aspect of behaviour or some kind of multivariate behavioural pattern.

The future behaviour of the user can then be compared with his profile in order to examine the consistency with it (normal behaviour) or any deviation from his profile, which may imply a fraudulent activity. An important issue is that one can never be certain that fraud has been perpetrated. Any analysis should only be treated as a method that provides us with an alert or a "suspicion score". That is, the analysis provides a measure that some observation is anomalous or more likely to be fraudulent than another. Special investigative attention should then be focused on those observations.

In telecommunications industry, user profiles can be constructed using appropriate usage characteristics. The aim is to distinguish a normal user from a fraudster. The latter is, in most of the cases, a user of the system who knows and mimics a normal user behaviour. All the data that can be used to monitor the usage of a telecommunication network are contained in the Call Detail Record (CDR) of any PBX. The CDR contains data such as: the caller ID, the chargeable duration of the call, the called party ID, the date and the time of the call, etc. In mobile telephone systems, such as GSM, the data records that contain details of every mobile phone attempt are the Toll Tickets.

When building a user profile, the first goal is to construct the basic building block that is a fundamental unit of comparison. Different units of comparison can be selected, depending of the type of the network and the type of fraud that is to be detected. One can use usage indicators related to the way a telephone is used, mobility indicators related to the mobility of the telephone if it is mobile and deductive indicators, which arise as a by-product of fraudulent behaviour, e.g. overlapping calls and velocity checks. The simplest usage indicator and the basic unit of comparison are the data per call, i.e. date and time, duration, caller ID, called No, and cost of call. Another simple unit can be a sequence of all the data of the calls that are made within a day. A third possible unit of comparison is the accumulated behaviour per day. That is, a sequence which is constructed by the number of calls made to local destinations, the duration (or the cost) of local calls, the number of calls to mobile destinations, the duration (or the cost) of mobile calls, the number of call to national or international destination and their corresponding duration. This per day accumulated behaviour of a user is a basic measure of the usage of his terminal and may be a measure that differentiates him from other users.

In Hilas and Sahalos (2007), an approach to user profiling in telecommunication was discussed, based on the latter basic unit of user behaviour. The empirical results demonstrate that such an approach yields high differentiation measures between users, and it is an interesting basis for future research. An important advantage

of this measure is that it hides all personal information of the user, e.g. caller or called party ID. This allows for the protection of the privacy of users during the experimentation for the development of any fraud detection technique.

3. Combining Fraud Detection Models

It is found that when a number of models are combined, instead of using a single model in isolation, there is improved performance. An example of combination of models is ensembling neural networks.

3.1 Ensembling Neural Networks Approaches

The neural network ensemble is a learning paradigm where a collection of a finite number of neural networks is trained for the same task (Zhou *et al.*, 2002). This shows that the generalization ability of a neural network system can be significantly improved through ensembling a number of neural networks, i.e. training many neural networks and then combining their predictions. In general, a neural network ensemble is constructed in two steps, i.e. training a number of component neural networks and then combining the component predictions.

As for the training component of neural networks, the most prevailing approaches are Bagging and Boosting (Zhou *et al.*, 2002). Bagging is proposed by Breiman based on bootstrap sampling. It generates several training sets from the original training set and then trains a component neural network from each of those training sets.

Boosting is proposed by Schapire and improved by Freund (Zhou *et al.*, 2002). It generates a series of component neural networks whose training sets are determined by the performance of former ones. Training instances that are wrongly predicted by former networks will play more important roles in the training of later networks. There are also many other approaches for training the component neural networks. Minjing (2006) utilises different object functions to train distinct component neural networks. Zhou *et al.* (2002) train component networks with different number of hidden units. Burge *et al.* (1997) initialize component networks at different points in the weight space.

As for combining the predictions of component neural networks, the most prevailing approaches are plurality voting or majority voting (Zhou *et al.*, 2002) for classification tasks, and simple averaging or weighted averaging for regression tasks. There are also many other approaches for combining predictions. However, in those approaches, the neural networks are in fact trained for different sub-tasks instead of the same task, which makes those approaches usually be categorized into mixture of experts instead of ensembles. Yet the goodness of such a process has not been formally proved. In Zhou *et al.* (2002) from the viewpoint of prediction, i.e. regression and classification, the relationship between the ensemble and its component neural networks was analysed, which revealed that ensembling many of the available neural networks might be better than ensembling all of those networks. Then, in order to show that those "many" neural networks can be effectively selected from a number of available neural networks, an approach named Genetic Algorithm based Selective Ensemble (GASEN) was presented. This approach selected some neural networks to constitute an ensemble according to some evolved weights that could characterize the fitness of including the networks in the ensemble. An empirical study on twenty big data sets shows that in most cases, the performance of the neural network ensembles generated by GASEN outperforms those generated by some popular ensemble approaches such as Bagging and Boosting in that GASEN utilizes far less component neural networks but achieves stronger generalization ability. Moreover, Zhou *et al.* (2002) employs the bias-variance decomposition to analyze the empirical results, which shows that the success of GASEN may owe to its ability of significantly reducing the bias along with the variance.

Random Rough Subspace method was proposed by Wei *et al*. (2011) and applied to detect anomaly in Insurance Industry. This method consists of several trained classifiers, the trained classifiers were combined using plural voting and output of the class was based on the outputs of these individual classifiers. The method employed by Wei *et al*. (2011) shows that the Random Rough Subspace method provides a faster and more accurate way to find suspicious insurance claims.

4. Bridging the Gap in Fraud Detection with Random Rough Subspace Based Neural Network Ensemble Method

As a result of the limitations of existing methods of fraud detection, the rough set ensemble method is proposed to bridge the gap. A random rough subspace based neural network ensemble method is employed in detecting subscription fraud in mobile telecommunication. This method involves creating a number of training subsets from the original training set. The training set is the CDR data consisting of the demography information of the customers of the network provider. A subset is created by randomly selecting samples from the original training set subspace. Each of the subsets is then used to train a neural network classifier; the trained neural network classifier are combined using the ensemble technique. The desired target function is approximated by averaging the classifiers. Using this technique, the probability of getting a prediction error will be very low compared to using a single classifier; thereby achieving a more accurate subscription fraud detection system.

For this study, four different training subsets were used to create four classifiers. A predicted target is obtained by averaging the outputs of the four classifiers. This helps to obtain an accurate prediction as long as at least three of the classifiers give accurate prediction with the absolute mean approximately equal to the target output given by the two classifiers. Figure 4 and Figure 5 show the block diagrams of the ANN ensemble.

The algorithm for the ensemble method is as shown below:

i. Given a training set S of size n, generate a new training set S1 of size n by sampling examples from S uniformly and with replacement,
ii. Repeat this sampling procedure, getting a sequence of four training sets,
iii. Run the learning algorithm four times, each time with a different training set,
iv. The ANN Ensemble Classifier then combines the predictions (t1, t2, t3, t4) of the individual classifiers to generate the final output t.

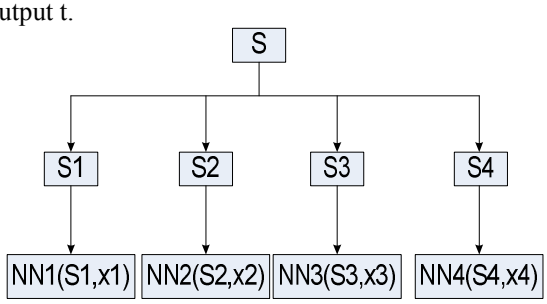

Figure 4. ANN Ensemble Training

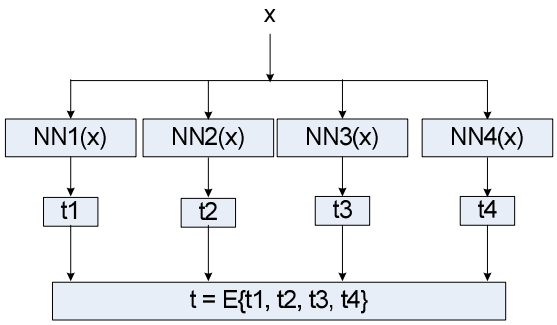

Figure 5. ANN Ensemble Testing

The subscription fraud detection system model is achieved by using sequences of call detail records (CDRs), which contain the details of each post-paid users on the network. The information produced for billing also contains usage behaviour information valuable for fraud detection. The proposed system model is presented in Figure 6.

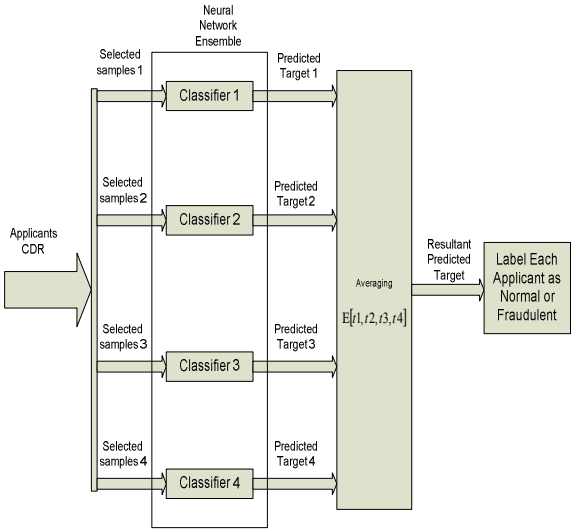

Figure 6. Rough Subspace Ensemble Subscription Fraud Detection Model

Fraud Classification

The CDR of 5120 customers is used for the model development. The variables TBI, LBA, MDC and DPR are used to classify each customer as Fraudulent or Normal. If a customer is detected to be fraudulent, the Phone blocked flag (PBF) is set to 1; while a non-fraudulent customer is assigned a PBF of 0.

Table 1 shows categorization of the subscribers while Figure 7 shows the DPR for the 5120 cases which are classified using the following rules and ordered according to their classes Fraud (1) and Normal (0):
IF (TBI>180) AND (LBA<500) THEN Fraud
IF (MDC>5000) AND (DPR >50) THEN Fraud
IF (TBI<=180) AND (LBA>500) THEN Normal
IF (MDC<=5000) AND (DPR <50) THEN Normal (Estevez *et al.*, 2005).

Table 1. Subscriber Categorization

Subscriber category	Number of cases
Normal	3994
Fraudulent	1126
TOTAL	**5120**

Training set

The ANN architecture consists of three (3) layers having:

i) 8 input neurons (national identity, address, age, income, phone number, gender, marital status, retire),
ii) 20 hidden neurons, and
iii) 1 output neuron.

The inputs used to train the ANN object are the CDR variables 1 to 8 while the target to the ANN is the phone blocked flag (PBF). The ANN object is trained to learn the characteristics of each customer's CDR as either Fraud or Normal. The learning process is repeated until the minimum error is obtained. The flow chart for the training stage of the ANN model is shown in Figure 7.

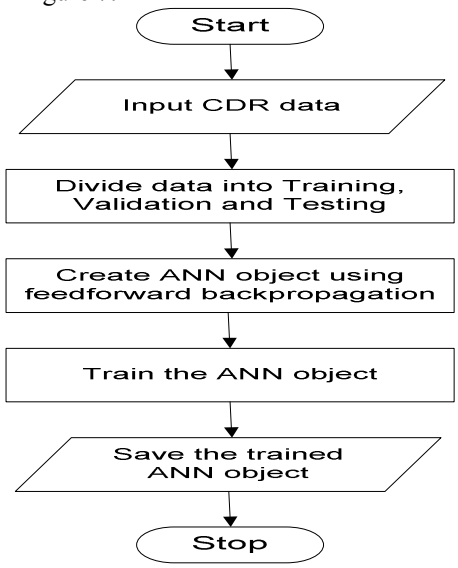

Figure 7. Flowchart of the Training Processes for the Neural Network

Fraud Prediction (Detection)

The subscription fraud detection is achieved by using customers' commercial antecedents which have been modeled in the ANN object. Some of the assumptions for predicting a fraudulent subscription are as follows:

- applicant's ID is similar to that of a fraudster,
- applicant's contact phone number is similar to a fraudster,
- applicant's address, age, gender and marital are similar to a fraudster.

A fraudster often uses the first line he used in committing fraud as the contact phone number during application for a new line. A particular number is supplied several times.

Testing Set

The developed ANN model was tested with different 2020 samples that are not part of the training set used to create the ANN model. Out of the testing samples, 1595 samples are fraudulent applications while 425 samples

are normal applications. The inputs to the ANN model are CDR variables 1 to 8. The model was used to simulate the inputs to predict the type of application (Fraud=1; Normal=0). The flow chart for the testing stage of the ANN model is shown in Figure 8.

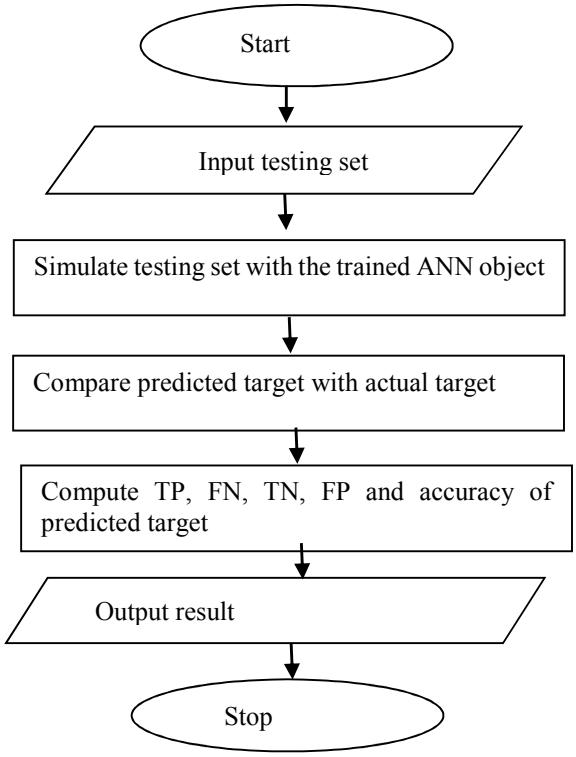

Figure 8. Testing Processes for the Neural Network

5. CONCLUSION

It has been shown that a combination of models is a powerful way to address the detection of fraud in telecommunications industry. This paper has addressed the detection of fraud in telecommunications industry by proposing a random rough subspace-based neural network ensemble method. The ensemble method consists of creation and random selection of training set from original set. Each subsets is used to train a neural network classifier in order to learn the characteristics of each customer. The desired target function is approximated by averaging the classifiers.

References

Burge, P., Shawe-Taylor, J., Cooke, C., Moreau, Y., Preneel, B., and Stoermann, C. (1997). Fraud Detection and Management in Mobile Telecommunication Networks. Royal Holloway University of London, England; Vodafone, England; ESAT.K.U. Leuven, Belgium; Siemens A.G. Germany.

Gopal, R. K. and Meher, S. K. (2007). A Rule-based Approach for Anomaly Detection in Subscriber Usage Pattern. *International Journal of Engineering and Applied Sciences* Vol. 3, pp 7.

Hilas, C. S. and Sahalos, J. N. (2007). User Profiling for Fraud Detection in Telecommunication Networks. *Technological Educational Institute of Serres, Serres 621 24, Greece. Aristotle University of Thessaloniki, Thessaloniki,* 541 24, Greece.

Minjing, P. (2006). A Neural Networks Ensemble Based Demand Forecasting Model for Third party Logistics. School of Business Administration South China University of Technology, Guangzhou, Guangdong, P.R.China, 510641. School of Management Wuyi University, Jiangmen, Guangdong, P.R.China, 529020.

Nelson, O. (2009). Detection of Subscription Fraud in Telecommunication Using Decision Tree Learning. Master's thesis, School of Graduate Studies, Makerere University.

Patidar, R. and Sharma, L. (2011). Credit Card Fraud Detection Using Neural Network. *International Journal of Soft Computing and Engineering (IJSCE)*, Vol 1, Issue-NCAI2011.

Rajani, S. and Padmavathamma, M. (2012). A Model for Rule Based Fraud Detection in Telecommunication. *International Journal of Engineering Research and Technology (IJERT),* Vol. 1, Issue 5, pp 1-7.

Rosas, E. and Analide, C. (2009). Telecommunication Fraud: Problem Analysis of an Agentbased KDD Perspective. *Department of Informatics University of Minho Braga, Portugal.*

Sanver, M. and Karahoca, A. (2008). Fraud Detection Using an Adaptive Neuro-Fuzzy Inference System in Mobile Telecommunication Networks. *Institute for Computational and Mathematical Engineering, Stanford University,*

Stanford, 94305, USA. Department of Computer Engineering, Bahcesehir University, Besiktas, Istanbul, 34900, Turkey.
Serrano, A. M. R., da Costa, J. P. C. L., Cardonha, C. H., Fernandes, A. A., and de Sousa Junior, R. T. (2010). Neural Network Predictor for Fraud Detection: A Study Case for the Federal Patrimony Department. *IBM Research Sao Paulo, Brazil,* (DOI: 10.5769/C2012010).
Taniguchi, M., Haft, M., Hollman, J., and Tresp, V. (2000). Fraud Detection in Communications Networks Using Neural and Probabilistic Methods. *Siemens AG, Corporate Technology Department Information and Communications* D-81730 Munich, Germany.
Verrelst, H., Lerouge, E., Moreau, Y., Vandewalle, J., Strmann, C., and Burge, P. (2001). A Rule Based and Neural Network System for Fraud Detection in Mobile Communications. Katholieke Universiteit Leuven, Belgium, Siemens Research, Mnchen, Germany, Royal Holloway University of London, UK.
Wei, X., Shengan, W., Dailing, Z. and Yang, B. (2011). Random Rough Subspace-based Neural Network Ensemble for Insurance Fraud Detection. *Fourth International Joint Conference on Computational Science and Optimization* Vol. 2, pp1276-1280.
Zhou, Z.-H., Wu, J., and Tang, W. (2002). Ensembling Neural Networks: Many Could Be Better Than All. *Artificial Intelligence Elsevier, 137(1-2):*239-263.

About the Authors

Awoyelu Iyabo holds a Ph.D. Degree in Computer Science from the Department of Computer Science and Engineering, Obafemi Awolowo University, Ile- Ife, Nigeria. At present, she is an academic staff of the Department of Computer Science and Engineering, Obafemi Awolowo University, Ile- Ife, Nigeria. She belongs to Data mining and Data warehousing Research Subgroup of Data Communications Research Group in the Department of Computer Science and Engineering, Obafemi Awolowo University, Ile- Ife, Nigeria.
Her research interests are in Data warehouse, Data mining and Recommender Systems.

Adenike Adebomi holds a B.Tech Degree in Computer Science from Ladoke Akintola University of Technology and M.Sc. Degree in Computer Science from the Department of Computer Science and Engineering, Obafemi Awolowo University, Ile- Ife, Nigeria. Her research interest is in Data mining.

Adekemi Amoo. holds a M.Sc. Degree in Computer Science from the Department of Computer Science and Engineering, Obafemi Awolowo University, Ile- Ife, Nigeria. At present, she is a Ph.D. student and an academic member of staff of the Department of Computer Science and Engineering, Obafemi Awolowo University, Ile- Ife, Nigeria. She belongs to Data mining and Data warehousing Research Subgroup of Data Communications Research Group in the Department of Computer Science and Engineering, Obafemi Awolowo University, Ile- Ife, Nigeria. Her research interests are in Data warehouse and Data mining.

Rachael Adebisi holds a B.Sc (Computer Science) of University of Ilorin, M.Ed. (Curriculum Studies), OAU and M.Sc. degree in Computer Science from the Department of Computer Science and Engineering, Obafemi Awolowo University, Ile-Ife, Nigeria. Her research areas are Expert Systems and Data mining.

Charles Mabude. holds a M.Sc. Degree in Computer Science from the Department of Computer Science and Engineering, Obafemi Awolowo University, Ile- Ife, Nigeria. His research interests are in Data mining and Recommender Systems.